On the sum of the digits of multiples

Antal Balog, Cécile Dartyge

On the sum of the digits of multiples

Antal Balog[*]
Alfréd Rényi Institute of Mathematics
Hungarian Academy of Sciences
H-1364 Budapest, P.O. Box: 127, Hungary
`balog@renyi.hu`

Cécile Dartyge[†]
Institut Élie Cartan
Université Henri Poincaré Nancy 1, BP 239
54506 Vandoeuvre Cedex, France
`Cecile.Dartyge@iecn.u-nancy.fr`

December 20, 2011

Abstract

Let $s_g(n)$ be the sum of digits of n when it is written on base g. Almost all integers n have a multiple hn with $s_g(hn)$ is significantly smaller than $s_g(n)$. We consider the opposite direction. For any fixed $K \geq 1$ we give a lower bound for the frequency of those integers n satisfying $s_g(hn) \geq \frac{1}{K}\, s_g(n)$ for all integers $h \geq 1$.

Mathematical Subject Classification (2010): 11A63, 11A25, 11K16

Keywords: radix representation, sum of digits

[*]Research was supported in part by Hungarian National Science Foundation Grants K72731 and K81658.

[†]Research was supported in part by grant ANR-BLAN 0103 MUNUM.

1 Introduction.

Let $g \geq 2$ be an integer and $s_g(n)$ denote the sum of digits in base g of the positive integer n. The relationship between $s_g(n)$ and $s_g(hn)$ raises interesting questions. We introduce the function

$$F(n) := \min_{h \geq 1} s_g(hn).$$

Clearly $F(n) \leq s_g(n)$. In fact in [1] it is proved that if n is not a power of g then there exists $h \notin \{g^k : k \in \mathbb{N}\}$ such that $s_g(hn) = s_g(n)$. On the other hand, $F(n)$ is much smaller than $s_g(n)$ in many cases, for example, if n is a prime and g is a quadratic non–residue modulo n, then by Euler lemma $g^{(n-1)/2} + 1 \equiv 0 \pmod{n}$, that is $F(n) = 2$. It is somewhat harder to find examples for $F(n)$ big. A natural question is to study the following set :

$$S_g(x) := \{n \leq x : F(n) = s_g(n)\},$$

and more generally, for any given $K \geq 1$,

$$S_g(x, K) := \{n \leq x : F(n) \geq \frac{1}{K} s_g(n)\}.$$

In [1] it is proved that these sets are small, more precisely :

$$|\{n \leq x : F(n) \geq \frac{1}{K} s_g(n)\}| \ll \frac{x}{(\log x)^{1/2}}.$$

In 1980, Stolarsky [4] obtained the following lower bound (strictly speaking he only studied the case $g = 2$, but his method worked in general)

$$|S_g(x)| \geq \frac{1}{g} x^{1/2}.$$

Our main objective in this paper is to improve the exponent in the lower bound and to extend the result to the sets $S_g(x, K)$.

Theorem 1. *For any fixed integers $g \geq 2$, $K \geq 1$ and $k \geq 0$ we have*

$$|S_g(x, K)| \geq \frac{1}{g^{K-1}\binom{g+k}{k+1}} x^{\theta},$$

with

$$\theta = \theta_{g,K,k} = \frac{(K-1)\log g + \log \binom{g+k}{k+1}}{(k+K+1)\log g}. \tag{1}$$

In the classical case of $K = 1$ Stolarsky's result corresponds to the choice of $k = 0$, while our method always gives a better exponent. For example, when $g = 2$ we take $k = 1$ to get

$$|S_2(x)| \geq \frac{1}{3}\, x^{\theta}, \quad \theta = \frac{\log 3}{3 \log 2} = 0.5283\ldots. \tag{2}$$

Another numerical examples are

g	10	100	1000	10^6	10^{12}	10^{24}
θ	$0.5856\ldots$	$0.6645\ldots$	$0.7213\ldots$	$0.8119\ldots$	$0.8808\ldots$	$0.9277\ldots$

Using the straightforward bounds

$$\frac{g^{k+1}}{(k+1)^k} \leq \binom{g+k}{k+1} \leq g^{k+1},$$

the choice $k \leq \log g < k + 1$ implies

Corollary 2. *Let $g \geq 2$ and $K \geq 1$ be integers. We have*

$$|S_g(x, K)| \geq \frac{1}{g^{K+\log g}}\, x^{1-\frac{1+\log(1+\log g)}{K+\log g}}.$$

Note that everything is explicit in Corollary 2, allowing us to choose K or g as functions of x. At the end of the paper we will indicate that the above choice of k provides the optimum of the present method, at least when g is big. The same calculation indicates that for $K \gg \log g$, the optimal choice is $k = 0$. In other words our iterative process does not improve upon the classical argument, which is

$$|S_g(x, K)| \geq \frac{1}{g^K}\, x^{1-\frac{1}{K+1}}.$$

A little twist on the iteration, however, improves this to

Theorem 3. *Let $g \geq 2$ and $K \geq 1$ be integers. We have*

$$|S_g(x, K)| \gg x^{1-\frac{1}{2K}}.$$

We did not make any effort to get the best multiplicative constants in these lower bounds. In Theorem 3 we did not even compute it explicitly, as it was too clumsy. On the other hand, Theorem 1 reflects our best exponent for $K = 1$, while Theorem 3 reflects our best exponent for $K \gg \log g$. A natural question is weather it is possible to combine te proofs of the two theorems. We will see after the proof of Theorem 3 that the exponent for (1)

$$\theta'_{g,K,k} = \frac{(2K-2)\log g + \log \binom{\lfloor \frac{g-1}{2} \rfloor + k + 1}{k+1}}{(2K+k)\log g} \tag{3}$$

can be derived easily by recent results of Mauduit, Pomerance and Sárközy [2] on the integers with a fixed sum of digits. This exponent is interesting only for g large enough. With more care, it is probably possible to improve the exponent (1) to

$$\theta_{g,K,k} = \frac{(2K-2)\log g + \log \binom{g+k}{k+1}}{(k+2K)\log g},$$

and this is superior for $3 \leq K \ll \log g / \log\log g$ but to obtain more significant improvement we need further ideas.

The function $s_g(n)$ is the ultimate example of a 'q–additive' function. Our arguments may be useful in other situations, and some of our statements may extend to a wider subclass of q–additive functions.

Acknowledgement. The authors are thankful to the referee for pointing out an error in the first version of this work and for other useful remarks.

2 Stolarsky's construction revisited.

Let us recall a classical fact about the sum of digit function.

s_g is sub-additive: for all $m, n \in \mathbb{N}$

$$s_g(m+n) \leq s_g(m) + s_g(n), \tag{4}$$

Formula (4) is easily obtained if m is of type $m = ug^k$ with $0 \leq u \leq g-1$. The general case for m is then obtained by iteration. It is also easy to see that there is equality in (4) if and only if "there is no carry in adding m and n". Equivalently, we have for $n > m$

$$s_g(n-m) \geq s_g(n) - s_g(m), \tag{5}$$

and there is equality in (5) if and only if no digits of m are bigger than the corresponding digits of n. This situation will be denoted by

$$n \trianglerighteq m.$$

The idea of Stolarsky was to consider the binary sum of digits of the multiples of $1 + 2^{\ell} + \cdots + 2^{\ell(r-1)} =: n$. He proved that $s_2(hn) \geq s_2(n)$ $\forall h \in \mathbb{N}$.

A natural extension of this idea to general basis is to consider the multiples of the integers

$$N_r := (g-1)(1 + g + \cdots + g^r) = g^{r+1} - 1,$$

with $r \geq 1$.

Lemma 4. *For all $h \geq 1$ we have*

$$s_g(hN_r) \geq s_g(N_r). \tag{6}$$

This lemma follows from the next by induction.

Lemma 5. *Writing $h = u + vg^{r+1}$, where $0 < u < g^{r+1}$ and $0 \leq v$ we have*

$$s_g(hN_r) \geq s_g((v+1)N_r). \tag{7}$$

For $1 \leq h \leq g^{r+1}$ we have

$$s_g(hN_r) = s_g(N_r). \tag{8}$$

Proof. Let $h > 1$. The obvious identity $(h-1)(N_r + 1) = (h-1)g^{r+1}$ implies

$$hN_r = N_r - (h-1) + (h-1)g^{r+1} = N_r - (u-1) + (h-1-v)g^{r+1}.$$

Note that $0 < N_r - (u-1) < g^{r+1}$ and that the least significant $r+1$ digits of $(h-1-v)g^{r+1}$ are zero. Thus we have:

$$s_g(hN_r) = s_g(N_r - (u-1)) + s_g(h-1-v). \tag{9}$$

If $v = 0$ then $h - 1 = u - 1$ and we have by (5) (since $N_r \trianglerighteq u - 1$) :

$$s_g(hN_r) = s_g(N_r) - s_g(u-1) + s_g(u-1) = s_g(N_r).$$

This proves the second statement (the case $h = g^{r+1}$ being trivial). If $v \geq 1$ then

$$h - 1 - v = u + vg^{r+1} - 1 - v = (v+1)N_r - (N_r - (u-1)) \tag{10}$$

and by (5) we find

$$s_g(h - 1 - v) \geq s_g((v+1)N_r) - s_g(N_r - (u-1)).$$

Inserting this into (9) gives

$$s_g(hN_r) \geq s_g((v+1)N_r).$$

This proves Lemma 5.

□

Proof of Lemma 4. The second statement of Lemma 5 proves (6) for $h \leq g^{r+1}$. Since $h = u + vg^{r+1} > v + 1$ we can use induction based on the first statement of Lemma 5 when $u \neq 0$, or based on $s_g(hN_r) = s_g(vN_r)$ when $u = 0$.

□

Let $\mathcal{H} := \{h \geq 1 : s_g(hN_r) = s_g(N_r)\}$. The importance of $\mathcal{H}$ is reflected in the fact that for any $h \in \mathcal{H}$ and for any $m \geq 1$ we have by Lemma 4

$$s_g(mhN_r) \geq s_g(N_r) = s_g(hN_r),$$

that is $hN_r \in S_g(x)$, whenever $hN_r \leq x$. We proved in Lemma 5 that $h \in \mathcal{H}$ if $1 \leq h \leq g^{r+1}$. This observation covers Stolarsky's argument.

3 How to go further?

For simpler exposition we introduce a variant of the relation $\trianglerighteq$. We write $n \trianglerighteq m \mod g^{r+1}$ to express that the least significant $r+1$ digits of n are not smaller than the corresponding digits of m. That is if the g–basis expansion of n and m are

$$n = n_0 + n_1 g + n_2 g^2 + \ldots, \quad m = m_0 + m_1 g + m_2 g^2 + \ldots,$$

then

$$n_j \geq m_j, \text{ for } j = 0, \ldots, r.$$

Lemma 6. *Let $h = u + vg^{r+1}$, where $1 \leq u < g^{r+1}$ and $0 \leq v$. If $v+1 \in \mathcal{H}$ and $u - 1 \trianglerighteq v \mod g^{r+1}$, then $h \in \mathcal{H}$.*

Proof. By (9) we have

$$s_g(hN_r) = s_g(N_r - (u-1)) + s_g(h - 1 - v).$$

By (10) we have

$$s_g(h - 1 - v) = s_g(u - 1 + vN_r) = s_g((v+1)N_r - (N_r - (u-1))).$$

If

$$(v+1)N_r \trianglerighteq N_r - (u-1), \tag{11}$$

then $s_g((v+1)N_r - (N_r - (u-1))) = s_g((v+1)N_r) - s_g(N_r - (u-1))$. In this case the condition $v+1 \in \mathcal{H}$ implies that $h \in \mathcal{H}$. Now it remains to check that (11) is satisfied. It is enough to compare the $r+1$ least significant digits as the right hand side itself is $< g^{r+1}$. Note the identity $(v+1)N_r = vg^{r+1} + N_r - v$. Since N_r is constructed to have maximal digits (11) is indeed equivalent to

$$N_r - v \trianglerighteq N_r - (u-1) \mod g^{r+1},$$

and this is equivalent to the condition $u - 1 \trianglerighteq v \mod g^{r+1}$. This ends the proof of Lemma 6.

□

Lemma 7. *Let $h \in \mathbb{N}$, $h = h_0 + h_1 g^{r+1} + \cdots + h_k g^{k(r+1)}$ with $0 \leq h_0, \ldots, h_k < g^{r+1}$. If $h_0 \geq 1$ and if $h_0 - 1 \trianglerighteq h_1 \trianglerighteq h_2 \trianglerighteq \cdots \trianglerighteq h_k$ then $h \in \mathcal{H}$.*

Proof. First note that the conditions imply

$$0 \leq h_k \leq \cdots \leq h_1 \leq h_0 - 1 < g^{r+1} - 1.$$

We prove this lemma by induction on k. The case $k = 0$ follows immediately from Lemma 5. Similarly for $k = 1$, $h = h_0 + h_1 g^{r+1}$, we have $h_1 + 1 \in \mathcal{H}$ by Lemma 5, and we can apply Lemma 6.

Suppose that the lemma is true for $k-1$ for some $k \geq 2$ and let $u = h_0$ and $v = h_1 + h_2 g^{r+1} + \cdots + h_k g^{(k-1)(r+1)}$. So that $h = u + vg^{r+1}$. By Lemma 6, $h \in \mathcal{H}$ whenever $u - 1 \trianglerighteq v \mod g^{r+1}$ and $v + 1 \in \mathcal{H}$. The first condition is clearly satisfied since $h_0 - 1 \trianglerighteq h_1$ and $v \equiv h_1 \mod g^{r+1}$. By our induction hypothesis the second condition also holds.

□

Lemma 8. *Let $\mathcal{E}_{k,r}$ be denote the set of the $k+1$ tuples $(h_0, h_1, \ldots, h_k)$ satisfying $0 \le h_0, \ldots, h_k < g^{r+1}$, $h_0 \ge 1$ and $h_0 - 1 \trianglerighteq h_1 \trianglerighteq \cdots \trianglerighteq h_k$. We have*

$$|\mathcal{E}_{k,r}| = \binom{g+k}{k+1}^{r+1} - \binom{g+k-1}{k}^{r+1}. \tag{12}$$

Proof. Note that $|\mathcal{E}_{0,r}| = g^{r+1} - 1$. Let $k \ge 1$. For all $(h_0, h_1, \ldots, h_k) \in \mathcal{E}_{k,r}$ we write

$$h_0 - 1 = \sum_{\ell=0}^{r} \varepsilon_0^{(\ell)} g^\ell,$$

$$h_j = \sum_{\ell=0}^{r} \varepsilon_j^{(\ell)} g^\ell, \quad j = 1, \ldots, k,$$

with $0 \le \varepsilon_j^{(\ell)} \le g - 1$.

The $k+1$–tuple $(h_0, \ldots, h_k)$ is in $\mathcal{E}_{k,r}$ if and only if the following conditions hold

$$0 \le h_0 - 1 < g^{r+1} - 1, \tag{13}$$

and for all $0 \le \ell \le r$

$$g - 1 \ge \varepsilon_0^{(\ell)} \ge \varepsilon_1^{(\ell)} \ge \varepsilon_2^{(\ell)} \ge \cdots \ge \varepsilon_k^{(\ell)} \ge 0. \tag{14}$$

For any given $0 \le \ell \le r$, the number of $(\varepsilon_0^{(\ell)}, \varepsilon_1^{(\ell)}, \ldots, \varepsilon_k^{(\ell)})$ satisfying (14) is $\binom{g+k}{k+1}$, the number of choices of $k+1$ not necessarily different elements out of $0, \ldots, g-1$, since each choice has exactly one decreasing order. Thus the number of all sets of $0 \le \varepsilon_j^{(\ell)} \le g-1$ satisfying (14) is $\binom{g+k}{k+1}^{r+1}$. However, sets with $\varepsilon_0^{(0)} = \cdots = \varepsilon_0^{(r)} = g - 1$, that is with $h_0 - 1 = g^{r+1} - 1$ $(= N_r)$, do not lead to a valid choice of h_0, as they violate (14). Much the same way, the number of $(\varepsilon_0^{(\ell)}, \varepsilon_1^{(\ell)}, \ldots, \varepsilon_k^{(\ell)})$ satisfying

$$g - 1 = \varepsilon_0^{(\ell)} \ge \varepsilon_1^{(\ell)} \ge \varepsilon_2^{(\ell)} \ge \cdots \ge \varepsilon_k^{(\ell)} \ge 0, \quad \ell = 0, \ldots, r$$

is $\binom{g+k-1}{k}^{r+1}$. This ends the proof of Lemma 8.

$\square$

4 The sets $S_g(x,K)$.

Now we consider the sets $\mathcal{H}_K := \{h \geq 1 : s_g(hN_r) \leq K s_g(N_r)\}$. Clearly $\mathcal{H}_1 = \mathcal{H}$. The importance of these sets is coded again in the fact, if $h \in \mathcal{H}_K$ then (by Lemma 4), for all $m \geq 1$ we have $s_g(mhN_r) \geq s_g(N_r) \geq \frac{1}{K} s_g(hN_r)$, that is $hN_r \in S_g(x,K)$ whenever $hN_r \leq x$.

Lemma 9. *Let $K \geq 2$ be an integer and $h = u + vg^{r+1}$, where $0 \leq u < g^{r+1}$. If $v \in \mathcal{H}_{K-1}$ then $h \in \mathcal{H}_K$.*

Proof. Here again we start out from the formula (9). There is nothing to prove for $u = 0$. For $u \geq 1$ we have

$$s_g(hN_r) = s_g(N_r - (u-1)) + s_g(h-1-v) = s_g(N_r - (u-1)) + s_g(u-1+vN_r).$$

Next we use the sub–additivity of s_g and the fact that N_r has maximal digits, that is $N_r \trianglerighteq u-1$

$$s_g(hN_r) \leq s_g(N_r - (u-1)) + s_g(u-1) + s_g(vN_r) = s_g(N_r) + s_g(vN_r).$$

Thus if $v \in \mathcal{H}_{K-1}$, then for all $0 \leq u < g^{r+1}$, $h = u + vg^{r+1} \in \mathcal{H}_K$, and we are done.

□

Proof of Theorem 1. In the previous section, actually in Lemma 7, we proved that $h = h_0 + h_1 g^{r+1} + \cdots + h_k g^{k(r+1)}$ with $0 \leq h_0, \ldots, h_k < g^{r+1}$ is in $\mathcal{H}$ whenever $h_0 \geq 1$ and $(h_0, \ldots, h_k) \in \mathcal{E}_{k,r}$. The same is true if $h_0 = 0$, $h_1 \geq 1$ and $(h_1, \ldots, h_k) \in \mathcal{E}_{k-1,r}$, and so on. All of these h satisfy $h < g^{(k+1)(r+1)}$. Adding also $h = g^{(k+1)(r+1)}$ we get

$$|\{h \in \mathcal{H} : h \leq g^{(k+1)(r+1)}\}| \geq |\mathcal{E}_{k,r}| + \cdots + |\mathcal{E}_{0,r}| + 1 = \binom{g+k}{k+1}^{r+1}.$$

Thus we have by Lemma 9

$$|\{h \in \mathcal{H}_2 : h \leq g^{(k+2)(r+1)}\}| \geq g^{r+1} \binom{g+k}{k+1}^{r+1}.$$

Iterating this observation we obtain

$$|\{h \in \mathcal{H}_K : h \leq g^{(k+K)(r+1)}\}| \geq g^{(K-1)(r+1)} \binom{g+k}{k+1}^{r+1}.$$

Finally we collect all these pieces. If $g^{(k+K+1)(r+1)} \le x$, $h \in \mathcal{H}_K$ and $h \le g^{(k+K)(r+1)}$, then $hN_r \in S_g(x,K)$, that is

$$|S_g(x,K)| \ge g^{(K-1)(r+1)} \binom{g+k}{k+1}^{r+1}. \tag{15}$$

For fixed g, K, k and (sufficiently large) x we choose the integer $r \ge 1$ by

$$g^{(k+K+1)(r+1)} \le x < g^{(k+K+1)(r+2)},$$

that is

$$r+2 > \frac{\log x}{(k+K+1)\log g}.$$

The lower bound in (15) can be modified as

$$|S_g(x,K)| \ge \frac{1}{g^{K-1}\binom{g+k}{k+1}} \left[g^{(K-1)} \binom{g+k}{k+1} \right]^{r+2} >$$

$$> \frac{1}{g^{K-1}\binom{g+k}{k+1}}\, x^{\theta},$$

with

$$\theta = \frac{(K-1)\log g + \log \binom{g+k}{k+1}}{(k+K+1)\log g},$$

as in (1). This completes the proof.

□

Proof of Theorem 3. We can suppose $K \ge 2$, as for $K=1$ this follows by taking $k=0$ into Theorem 1. We suppose that h is of the following type

$$h = u + v_1 g^{r+1} + \cdots + v_{K-1} g^{(K-1)(r+1)} + w_1 g^{K(r+1)} + \cdots + w_{K-1} g^{(2K-2)(r+1)},$$

where

$$\begin{aligned} 1 \le u \le v_1 \le \cdots \le v_{K-1} < g^{r+1}, \\ 0 \le w_{K-1} \le \cdots \le w_1 \le v_{K-1} - 1. \end{aligned} \tag{16}$$

We start from the identity

$$hN_r =$$
$$= \big(N_r - (u-1)\big) + \big(N_r - (v_1 - u)\big) g^{r+1} + \cdots + \big(N_r - (v_{K-1} - v_{K-2})\big) g^{(K-1)(r+1)} +$$
$$+ \big(v_{K-1} - w_1 - 1\big) g^{K(r+1)} + \big(w_1 - w_2\big) g^{(K+1)(r+1)} + \cdots + w_{K-1} g^{(2K-1)(r+1)}.$$

In the above representation all coefficients are non negative and $< g^{r+1}$, moreover $N_r \trianglerighteq u-1, \dots, N_r \trianglerighteq v_{K-1}-v_{K-2}$, and we conclude

$$s_g(hN_r) = Ks_g(N_r) - s_g(u-1) - s_g(v_1-u) - \cdots - s_g(v_{K-1}-v_{K-2}) +$$

$$+ s_g(v_{K-1}-w_1-1) + s_g(w_1-w_2) + \cdots + s_g(w_{K-1}).$$

Thus $h \in \mathcal{H}_K$ whenever

$$\begin{aligned} &s_g(v_{K-1}-w_1-1) + s_g(w_1-w_2) + \cdots + s_g(w_{K-1}) \le \\ &\le s_g(u-1) + s_g(v_1-u) + \cdots + s_g(v_{K-1}-v_{K-2}). \end{aligned} \tag{17}$$

There is a one–to–one correspondence between systems of integers satisfying (16) and systems of integers satisfying

$$0 \le X_1, \dots, X_K, Y_1, \dots, Y_K, \quad X_1 + \cdots + X_K = Y_1 + \cdots + Y_K < N_r. \tag{18}$$

The correspondence is given by

$$\begin{array}{llll} u & = 1 + Y_1 & w_1 & = X_2 + \cdots + X_K \\ v_1 & = 1 + Y_1 + Y_2 & \vdots & \\ \vdots & & w_{K-2} & = X_{K-1} + X_K \\ v_{K-1} & = 1 + Y_1 + \cdots + Y_K & w_{K-1} & = X_K \end{array}$$

and the condition (17) is equivalent to

$$s_g(X_1) + \cdots + s_g(X_K) \le s_g(Y_1) + \cdots + s_g(Y_K). \tag{19}$$

It is an easy classical problem to compute M, the number of solutions to (18). If a solution of (18) does not satisfy (19), then changing the role of $(X_1, \dots, X_K)$ and $(Y_1, \dots, Y_K)$ we get one that satisfies (19). In other words, at least half of the solutions of (18) satisfy (19) as well. To compute M we write

$$r(n) = |\{n = X_1 + \cdots + X_K,\ 0 \le X_j\}| = \binom{n+K-1}{K-1}.$$

Finally we have

$$M = \sum_{n=0}^{N_r-1} \big(r(n)\big)^2 \ge \sum_{n=0}^{N_r-1} \binom{n+K-1}{K-1}^2 \ge \frac{N_r^{2K-1}}{(2K-1)((K-1)!)^2}.$$

We constructed at least $M/2$ elements h of $\mathcal{H}_K$, satisfying $h < g^{(2K-1)(r+1)}$. With these h we have $hN_r \in S_g(x,K)$ whenever $hN_r \le x$ which is the case if $g^{2K(r+1)} \le x < g^{2K(r+2)}$. Collecting all pieces we arrive at

$$|S_g(x,K)| \ge \frac{N_r^{2K-1}}{2(2K-1)((K-1)!)^2} \gg g^{(2K-1)(r+1)} \gg x^{1-\frac{1}{2K}}.$$

$\square$

Outlines of the proof of (3). We will present now the main ideas of the proof of (3). Since the improvement is not important and interesting only for large g, we won't write all the details. A natural way to combine the two previous proofs is to considere integers h of type:

$$\begin{aligned} h &= u + v_1 g^{r+1} + \cdots + v_{K-1} g^{(K-1)(r+1)} + w_1 g^{K(r+1)} + \cdots + w_{K-1} g^{(2K-2)(r+1)} \\ &+ h_1 g^{(2K-1)(r+1)} + \cdots + h_k g^{(2K-2+k)(r+1)}, \end{aligned}$$

where $u, v_1, \ldots, v_{K-1}, w_1, \ldots, w_{K-1}$ satisfy (16) and

$$w_{K-1} \trianglerighteq h_1 \trianglerighteq \cdots \trianglerighteq h_k. \tag{20}$$

We will also suppose that all the digits of w_{K-1} are less than $\left\lfloor \frac{g-1}{2} \right\rfloor$:

$$w_{K-1} = \sum_{j=0}^{r} \varepsilon_j g^j \quad \text{with } 0 \le \varepsilon_0, \ldots, \varepsilon_r \le \left\lfloor \frac{g-1}{2} \right\rfloor. \tag{21}$$

By the previous computations we remark that $h \in \mathcal{H}_K$ whenever (17) holds.

Mauduit, Pomerance and Sárközy ([2] Lemma 4) proved uniformly for $\lambda > 0$ and $N > N_0(g)$:

$$\sum_{\substack{n \le N \\ \left| s_g(n) - \frac{(g-1)}{2} \left\lfloor \frac{\log N}{\log g} \right\rfloor \right| > \lambda \sqrt{\left\lfloor \frac{\log N}{\log g} \right\rfloor}}} 1 \le c_1(g) \max\left(N \exp\left(\frac{-6\lambda^2}{g^2-1} \right), N^{1-\frac{c_2(g)}{\log\log N}} \right), \tag{22}$$

with $c_1(g) > 0$, $c_2(g) > 0$. By Lemma 5 of [2] we also have for any integer μ

$$\sum_{\substack{n \le N \\ s_g(n) = \frac{(g-1)}{2} \left\lfloor \frac{\log N}{\log g} \right\rfloor + \mu \lfloor \sqrt{\log N} \rfloor}} 1 \gg_{g,\mu} \frac{N}{\sqrt{\log N}}. \tag{23}$$

Their results provide in fact asymptotic formulae valid in a large range for the parameters (see also [3]).

Applying $2K-2$ times (22) for some $\lambda = \lambda(g,K)$ large enough and using also (23), we can prove that for any w_{K-1} verifying (21) there exists $C(g,K) > 0$ such that there are at least $C(g,K)g^{(2K-2)(r+1)}/\sqrt{r}$ integers $u, v_1, \ldots, v_{K-1}, w_1, \ldots, w_{K-2}$ satisfying the following inequalities :

$$s_g(u-1) \geq \frac{(g-1)(r+1)}{2} + 2K\lambda\sqrt{r},$$

$$\max(s_g(v_{K-1}-w_1-1), s_g(w_1-w_2), \ldots, s_g(w_{K-2}-w_{K-1}) \leq \frac{(g-1)(r+1)}{2} + \lambda\sqrt{r},$$

$$\min(s_g(v_1-u), s_g(v_2-v_1), \ldots, s_g(v_{K-1}-v_{K-2})) \geq \frac{(g-1)(r+1)}{2} - \lambda\sqrt{r}.$$

Thus (17) holds for these integers $u, v_1, \ldots, v_{K-1}, w_1, \ldots, w_{K-1}$. To end the proof of (3) it remains to see that the number of $w_{K-1}, h_1, \ldots, h_k$ satisfying (21) and (20) is $\binom{\lfloor \frac{g-1}{2} \rfloor + k + 1}{k+1}^{r+1}$.

5 Some computations on the θ.

In this last section we justify the choice of k in the cases mentioned in the Introduction. For fixed g and K we denote $\theta_k = \theta_{g,K,k}$, see (1). We will be brief with some details of computations. Since

$$\theta_k - \theta_{k+1} = \frac{\theta_{k+1}\log g - \log \frac{g+k+1}{k+2}}{(k+K+1)\log g},$$

we have for all $k \geq 0$ that

$$\theta_k \geq \theta_{k+1} \quad \text{if and only if} \quad \theta_{k+1}\log g \geq \log \frac{g+k+1}{k+2}. \tag{24}$$

In other words, a sufficiently strong lower bound for θ_{k+1} for all $k \geq k_0$ implies that the optimal choice of k in Theorem 1 satisfies $k \leq k_0$. Indeed, we prove

Lemma 10. *If $K \gg \log g$ then*

$$\theta_k \geq \theta_{k+1} \quad \textit{for all} \quad k \geq 0.$$

For all $K \geq 1$ we have

$$\theta_k \geq \theta_{k+1} \quad \textit{if} \quad k \gg \log g.$$

Proof. Since $(g+t)/(t+1)$ is a decreasing function of the variable t on the positive real numbers, we have

$$(k+K+2)\,\theta_{k+1}\log g = (K-1)\log g + \sum_{j=0}^{k+1}\log\frac{g+j}{j+1} \geq$$

$$\geq K\log g + \log\frac{g+k+1}{k+2} + \int_1^{k+1}\log\frac{g+t}{t+1}\,dt.$$

Computing the integral we get from (24) that $\theta_k \geq \theta_{k+1}$ follows if

$$(K-1)\left(\log(k+2) - \log\frac{g+k+1}{g}\right) + (g-1)\log\frac{g+k+1}{g+1} + 2\log 2 \geq$$

$$\geq 2\log(g+1) - \log g,$$

and this follows if

$$(K-1)\log\frac{4}{3} + (g-1)\log\frac{g+k+1}{g+1} \geq \log(g+1), \tag{25}$$

by the inequalities

$$\log(k+2) - \log\frac{g+k+1}{g} \geq \log\frac{4}{3} \quad \text{for all} \quad g\geq 2, k\geq 0,$$

and

$$2\log(g+1) - \log g - 2\log 2 \leq \log(g+1) \quad \text{for all} \quad g \geq 2.$$

Finally one can see that for $K \gg \log g$ the first term, for $k \gg \log g$ the second term in the left hand side of (25) majorize the right hand side. This proves the lemma. □

References

[1] C. Dartyge, F. Luca and P. Stănică, On digit sums of multiples of an integer, Journal of Number Theory 129, (2009), 2820-2830.

[2] C. Mauduit, C. Pomerance and A. Sárközy, On the distribution in residue classes of integers with a fixed digit sum', The Ramanujan J. 9 (2005), 45-62.

[3] C. Mauduit and A. Sárközy, On the arithmetic structure of the integers whose sum of digits is fixed, Acta Arith. LXXXI.2 (1997), 145-173.

[4] K. B. Stolarsky, Integers whose multiples have anomalous digital frequencies, *Acta Arith.* **38** (1980), 117–128.